

# Série Simple

## caractère

By: GOURAI KAWTAR

### Quantitatif

discret: dénombrable  
- accepte la répétition  
- fini

continu: infini  
- n'accepte pas la répétition (prévalence)

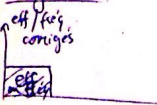
### qualitatif

#### discret (ssvi)

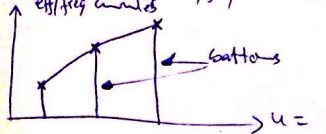
| val   | eff   | eff cum           |
|-------|-------|-------------------|
| $x_1$ | $n_1$ | $N_1 = n_1$       |
| $x_2$ | $n_2$ | $N_2 = n_1 + n_2$ |
| $x_m$ | $n_m$ | $N_m = n$         |

| freq  | freq cum          |
|-------|-------------------|
| $f_1$ | $F_1 = f_1$       |
| $f_2$ | $F_2 = f_1 + f_2$ |
| $f_m$ | $F_m = n$         |

#### Histogramme



Courbe des effs cumulés / freq. cumulés



#### continu (ssvc)

| classe           | eff   | eff cum                   | amplitude       |
|------------------|-------|---------------------------|-----------------|
| $[x_{i-1}, x_i]$ | $n_i$ | $N_i = n_1 + \dots + n_i$ | $x_i - x_{i-1}$ |
| $[x_{m-1}, x_m]$ | $n_m$ | $N_m = n$                 | $x_m - x_{m-1}$ |

Si les classes n'ont pas la même longueur (= amplitude)  
on effectue des corrections  
sur les effs ou les freq.

| classe           | eff   | eff cum | amplitude       | classe à représenter |
|------------------|-------|---------|-----------------|----------------------|
| $[x_{i-1}, x_i]$ | $n_i$ | $N_i$   | $x_i - x_{i-1}$ | $u = \dots$          |

| eff corrigé                           |
|---------------------------------------|
| $\frac{\text{eff}}{\text{amplitude}}$ |

## Catálogos des valeurs centrales classiques

### discret (ssvi)

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i y_i \text{ (calculatrice)}$$

Moyenne arithmétique → c'est une valeur algébrique  
n ne s'interprète pas seul

### Médiane

α = 1/2  
partage la série en 2 s. séries  
Selon la parité de m

$m \times \alpha \notin \mathbb{N}$  ou  $m = 2k+1$   
⇒  $m_e = x'_{[ \frac{m}{2} ] + 1}$   
partie entière

$m \times \alpha \in \mathbb{N}$  ou  $m = 2k$   
intervalle médian  
 $(x'_k, x'_{k+1}) = (x'_{\frac{m}{2}}, x'_{\frac{m}{2}+1})$

### continu (ssvc)

$$\bar{x} \approx \frac{1}{n} \sum_{i=1}^s n_i \theta_i$$

Moyenne arithmétique  
 $\theta_i = \frac{z_{i-1} + z_i}{2}$   
n ne s'interprète pas seul

### Médiane

à partir de la courbe des eff/freq cumulés  
 $\frac{m}{2} \in [N_{i-1}, N_i[$   
⇒  $m_e \in [z_{i-1}, z_i[$   
 $m_e \approx z_{i-1} + (z_i - z_{i-1}) \times \frac{\frac{m}{2} - N_{i-1}}{N_i - N_{i-1}}$   
ou bien  
 $m_e \approx z_{i-1} + (z_i - z_{i-1}) \times \frac{\frac{1}{2} - F_{i-1}}{F_i - F_{i-1}}$

### Quantile

$m \times \alpha \notin \mathbb{N}$   
 $Q_\alpha(u) = x'_{[ \frac{m}{\alpha} ] + 1}$   
prend  $Q_\alpha(u) = y_i$   
intervalle quantile décodé  
 $(x'_{i-1}, x'_{i+1})$   
en cherche un intervalle  $[F_{i-1}, F_i]$  qui corresponde à l'intervalle  $(y_{i-1}, y_i)$

### Quantile

$m \times \alpha \in [N_{i-1}, N_i[$   
⇒  $Q_\alpha(u) \in [z_{i-1}, z_i[$   
⇒  $Q_\alpha(u) \approx z_{i-1} + (z_i - z_{i-1}) \times \frac{\alpha - F_{i-1}}{F_i - F_{i-1}}$   
ou bien  
 $Q_\alpha(u) \approx z_{i-1} + (z_i - z_{i-1}) \times \frac{\frac{m}{\alpha} - N_{i-1}}{N_i - N_{i-1}}$

## Pr les caractères quantitatifs continus :

→ La masse :

$$m_i = \sum_{j: z_{i-1} \leq x_j < z_i} n_j = m_i \times \theta_i = m_i \times \frac{z_i + z_{i-1}}{2}$$

→ La médiane : c'est l'elt de la série la partage en 2 s.s de m masse totale moitié.

$$M_{le} = \frac{M_t}{2}$$

$$\frac{M_t}{2} \in [M_{i-1}, M_i] \Rightarrow M_{le} \in [z_{i-1}, z_i]$$

$$M_{le} \approx z_{i-1} + (z_i - z_{i-1}) \times \frac{\frac{M_t}{2} - M_{i-1}}{M_i - M_{i-1}}$$

NB :  $m_e < M_{le}$  (si  $M_{le} > \frac{m}{2} N_{me}$  ou  $M_{me} < \frac{M_t}{2}$ )

→ Indice de concentration :

$$i_c \approx 1 - \sum_{i=0}^s f_i (R_{i-1} + R_i)$$

By: GOURAI KAWTAR

Si  $0 < i_c < 0,2 \Rightarrow$  série à peu près égalitaire

Si non  $\Rightarrow$  série s'éloigne d'une série égalitaire.



Si on nous demande de ~~fait~~ d'étudier  
la dispersion des élt's de la série p/p à la moyenne  
on calcule  $f_n$  + q  $f_n = \frac{S_n \rightarrow \text{écart-type de } n}{|\bar{x}|}$

Si  $f_n < 3 \Rightarrow$  faible dispersion p/p à  $\bar{x}$

Si  $f_n > 3 \Rightarrow$  forte dispersion p/p à  $\bar{x}$ .

By: GOURAT KAWTAR

Pour tracer la courbe des concentrations

on calcule  $F_n$  et  $R_n$  + q

$$F_n = \frac{N_n}{n} = \frac{\text{eff. cumu}}{\text{eff. total}} \quad \left( \begin{array}{l} \text{si on veut } F_n \text{ en \%} \\ F_n \times 100 \end{array} \right)$$

$$R_n = \frac{M_n}{M_t} = \frac{\text{Masse cumu}}{\text{Masse totale}} \quad \left( \begin{array}{l} \text{si on veut } R_n \text{ en \%} \\ R_n \times 100 \end{array} \right)$$

Règle de STURGE :

$$\text{Nb're des classe} = \frac{\text{étendue}}{\text{la longueur d'1 classe}}$$

$$\text{étendue} = (\sup u_i) - (\inf u_i)$$



# Série double

(cas d'un cara conjoint quantitatif)

By: GOURAT KAWTAR

Mise en ordre

$$\Rightarrow (x_1', y_1'), \dots, (x_i', y_s')$$

→ Tableau de correspondances (tableau des eff.)

→ " de contingences ( " des fréquences)

Déduire

⇒

1

Séries marginales

↳ obtenue en faisant la somme des effectifs ou des fréq. des lignes pr X ou des colonnes pr Y

de X:  $(x_i', n_{i \cdot} (\text{ou } f_{i \cdot}))_{i=1, \dots, r}$

de Y:  $(y_j', n_{\cdot j} (\text{ou } f_{\cdot j}))_{j=1, \dots, s}$

Séries conditionnelles

↳ obtenue en fixant une colonne du tableau pr X ou une ligne pr Y

de X =  $x_i'$  fixé et Y variable:  $y_1', \dots, y_s'$

$Y/X = x_i'$

de Y =  $y_j'$  fixé et X variable:  $x_1', \dots, x_r'$

$X/Y = y_j'$

Covariance de la série double:

$$S_{xy} = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$$

$$S_{xy} = \sum xy \div n - \bar{x}\bar{y}$$

2

Etude de corrélation

Coeff. de corrélation linéaire:

$$r_{xy} = \frac{S_{xy}}{S_x \cdot S_y}$$

;  $r_{xy}$  mbre réel.

= 0

pas de corrélat°

≥ 0,1

fortement corrélés  
X et Y varient en m temps.

= 1

corrélation parfaite

dans ces cas, on peut remplacer y par ax + b sans commettre trop d'erreurs.

By: GOURAI KAWTAR

si  $|r_{xy}| > 0$

## Etuale de la régression :

on propose un modèle explicatif de  $y$  à partir de  $x$

$$y = f(x) + \varepsilon$$

← fct d'erreur du modèle

fct de régr. du modèle (à déterminer)

permet de déterminer  $a$  et  $b$ .

Principe des moindres carrés

(remplacer chaque valeur  $y_i$  de la s. double par  $f(x_i) + \varepsilon_i$ )

$$y_i = f(x_i) + \varepsilon_i \quad + q \quad \sum_{i=1}^n \varepsilon_i^2 = \sum_{i=1}^n (y_i - f(x_i))^2$$

Soit minimale.

on distingue 2 modèles

linéaire

de régr. exponentiel

$$f(x_i) = ax_i + b$$

$$y_i = a e^{bx_i}$$

$$\hat{a} = \frac{S_{xy}}{S_x^2}$$

$$\tilde{y}_i = \ln(y_i) = \ln(a) + bx_i$$

$$\hat{b} = \bar{y} - \frac{S_{xy}}{S_x^2} \cdot \bar{x}$$

$$\hat{b} = \frac{S_{\ln y}}{S_x^2}$$

droite de régr. de  $y$  en  $x$ :

$$\bar{y} - \bar{y} = \frac{S_{xy}}{S_x^2} (x - \bar{x})$$

$$\ln(\hat{a}) = \ln(\bar{y}) - \hat{b} \bar{x}$$

droite d'ajustement de  $y$  par  $x$  (ou de régression)

$$y = ax + b$$

$$a = \frac{S_{xy}}{S_x^2}$$

$$b = \bar{y} - \frac{S_{xy}}{S_x^2} \cdot \bar{x}$$

droite d'ajustement de  $x$  par  $y$  (ou de régression)

$$x = ay + b$$

$$a = \frac{S_{xy}}{S_y^2}$$

$$b = \bar{x} - \frac{S_{xy}}{S_y^2} \cdot \bar{y}$$